ProFORMA: Probabilistic Feature-based On-line Rapid Model Acquisition

Qi Pan
http://mi.eng.cam.ac.uk/~qp202
Gerhard Reitmayr
http://mi.eng.cam.ac.uk/~gr281
Tom Drummond
http://mi.eng.cam.ac.uk/~twd20

Machine Intelligence Laboratories
Department of Engineering
Cambridge University
Cambridge
United Kingdom

3D models have an extensive range of uses in computer vision, but generating these models can be a difficult and time consuming task. Whilst many automated and semi-automated 3D reconstruction techniques exist, current methods have a variety of drawbacks meaning that ultimately many high quality 3D models are still generated using labour intensive methods or completely by hand.

Off-line reconstruction methods have the potential to create highly detailed and accurate 3D reconstructions, but this level of detail usually implies high computational cost. This means that the reconstruction pipeline must be split into two phases: an image sequence or video collection phase (fast), and an off-line processing phase (slow), after which the model is obtained. If at this stage the model is found to be flawed, additional images or an entire video sequence must be collected and the lengthy reconstruction phase repeated. It may take several iterations and many hours before obtaining an acceptable model.

Our system, ProFORMA, enables textured 3D models to be acquired on-line in just over a minute, with models being reconstructed fast enough to provide feedback for view planning. The user is free to interact with the object which is robustly tracked using a method that makes immediate use of the partial model [2]. The reconstruction environment is not constrained to a studio and enables models of textured objects to be acquired using a single camera and commodity hardware. Rapid reconstruction is enabled by a novel recursive probabilistic tetrahedron carving algorithm, which uses visibility of point features to rapidly generate a surface model. Parts of an object not yet seen by the system are indicated to the user who can then manipulate the object in order to provide new views.

Our on-line reconstruction system uses a fixed-position video camera and allows the user to manipulate the object to be modelled using their hand. This enables the object to be orientated so that all parts of the object (including the base) can be viewed and modelled. No assumptions about the object are made and no prior information is known about the object, although the object must be sufficiently textured.

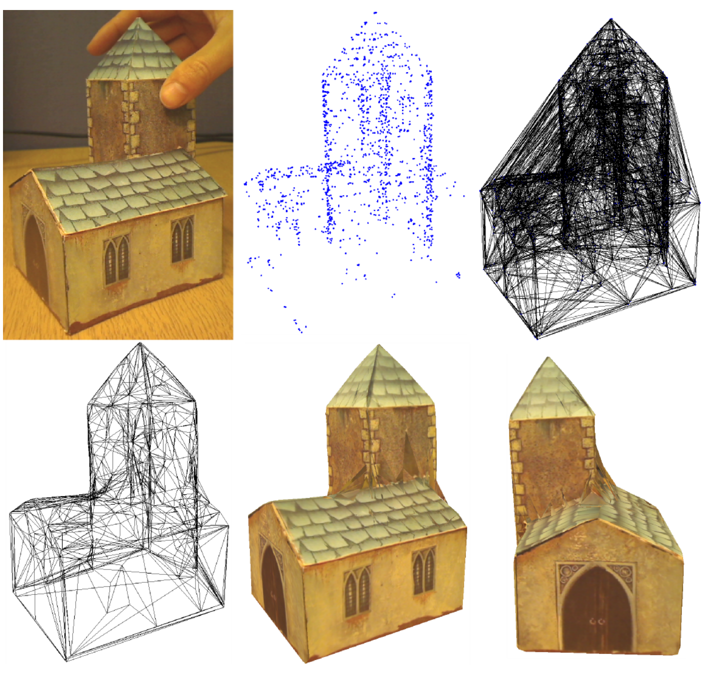

Figure 1: Results from each stage of the reconstruction pipeline. Left to right: (a) Object (b) Point cloud (c) Tetrahedralisation (d) Surface mesh (e) and (f) Rendered textured model

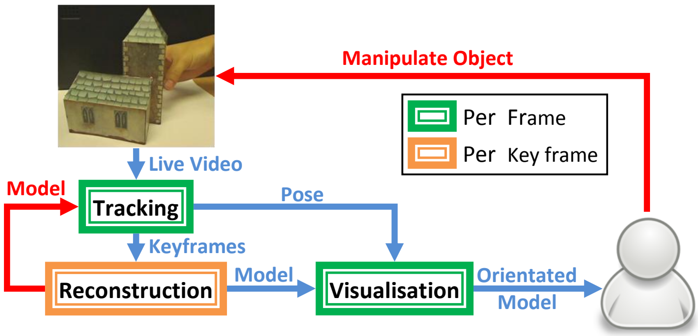

Figure 2: Flowchart of system operation.

Figure 2 shows the design of the system. The *tracking* thread takes a video input of the scene and calculates the pose of the camera relative to the object at frame rate. It also tracks the location of 2D image features which adhere to rigid body motion constraints (epipolar constraints). When a large enough rotation is detected, a keyframe is taken and passed to the *reconstruction* thread, which generates a partial 3D model of the object. This model is fed back into the tracker to provide additional structural information that can be used during tracking. The model is also utilised by the *visualisation* system to show the user the state of the reconstruction so far, with the model pose updated to the live pose of the object obtained from the tracking thread. The user provides feedback to the system by looking at the visualisation and manipulating the object accordingly to provide new views.

The reconstruction module processes keyframe information in a background thread, like Klein and Murray [1], allowing tracking to run in parallel. Keyframe information consists of 2D feature tracks and a list of 3D landmarks. Landmark positions are obtained by bundle adjustment of triangulated 2D feature tracks to obtain a point cloud (Figure 1). This only produces a very sparse model and needs to be processed in order to generate a surface model. A Delaunay tetrahedralisation is performed on the point cloud to obtain a partitioned convex hull. Tetrahedra are carved away based on landmark visibility to obtain the surface mesh.

Carving is performed probabilistically to reduce noise and performed recursively over only surface tetrahedra for efficiency. We draw rays emanating from the camera position at each keyframe to landmarks visible in that keyframe. These rays are intersected with triangles making up the Delaunay tedrahedralisation. Landmarks are viewed as observations of a surface triangle corrupted with Gaussian noise. We assume that if a triangle forms part of the surface mesh, landmarks corresponding to intersecting rays should exhibit Gaussian noise. This can be used this to calculate the probability of a triangle forming part of the surface mesh. Triangles (and their corresponding tetrahedra) are carved away if the probability is below a threshold to yield the surface model. This can then be textured mapped and displayed to the user to indicate the state of reconstruction.

ProFORMA enables complete textured models of objects to be acquired in just over a minute whilst also providing useful visual feedback for view planning. Reconstructed models can be used to robustly track the object even under large motions. The collected video sequence, confirmed to contain good quality views of the whole object, can then be passed to an offline reconstruction method to generate a higher quality model.

[1] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *Proc. Int'l Symposium on Mixed and Augmented Reality*, 2007.

[2] E. Rosten and T. Drummond. Fusing points and lines for high performance tracking. In *Proc. Int'l Conf. on Computer Vision*, 2005.